Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-02 | June-2024

Distinguishing Between Drones and Birds Using CNNs Algorithm

Zahraa Mohammed Sahib[1], Akmam Majed Mosa[2], Zahraa Hussein Ali[3], Dr Hayder A Nahi[4]
[1,2,3,4] Computer Center, Al-Qasim Green University, Babylon-Iraq.
comzahraa@gmail.com, akmammajed@uoqasim.edu.iq, zahraa2013.1988@gmail.com

DOI: 10.5281/zenodo.13920101

ABSTRACT

Recognizing drones or unmanned aerial vehicles (UAVs) from birds is a crucial capability for numerous applications. We create a convolutional neural network (CNN) drone identification system that can distinguish between images of drones and birds. A dataset of photos taken of birds and drones in various settings is used to train the CNN model. Our model distinguishes between drones and birds with 93% accuracy. The excellent results show that CNNs are capable of accurately differentiating between drones and birds under practical circumstances. Overall, this work demonstrates that deep learning may be used to achieve accurate drone recognition when similar avian items are present.

Keywords: Drone Detection, Bird Detection, Convolutional Neural Networks, Object Classification

Cite as: Zahraa Mohammed Sahib, Akmam Majed Mosa, Zahraa Hussein Ali, & Dr Hayder A Nahi. (2024). Distinguishing Between Drones and Birds Using CNNs Algorithm. *LC International Journal of STEM*, *5*(2), 1–10. **https://doi.org/10.5281/zenodo.13920101**

INTRODUCTION

Unmanned aerial vehicles (UAVs), sometimes known as drones, have seen a sharp rise in military, commercial, and recreational use in recent years. However, whether fitted with cameras or other payloads, the widespread use of drones poses serious security hazards. Consequently, it is now essential to detect and identify drones with reliability.

It can be difficult to distinguish drones from birds in the visual sense, especially when viewing them from a distance. Drones may also maneuver dynamically and fly swiftly, which makes detection more difficult. Due to variations in appearance, illumination, and perspective, it is still challenging to detect drones in open spaces. Lung. Neural networks (CNNs) are one kind of deep learning technology that has shown promise in identifying challenging items and classifying them. CNNS is able to recognize various kinds of objects from pictures by detecting strong patterns. [2].

In this work, we create a CNN-based drone recognition system to distinguish drone photos from bird photos. A wide variety of real-world drone and bird photo datasets are used to train the CNN model. Our tests show that the CNN can distinguish between drones and birds with 93% accuracy. The excellent result demonstrates CNNs' dependability in distinguishing between drones and related avian objects.

The layout of the paper is as follows: An overview of relevant drone detection research is given in Section II. Our CNN-based method for differentiating between drones and birds is explained in Section III. The experiment's findings are examined in Section IV. Section V concludes with the conclusions.

RELATED WORK

Being able to distinguish drones from birds is crucial for a variety of applications. Can telepathic neural networks (CNNs) are used.

Lin et al. devised a method for distinguishing between drones and birds in video data [3]. Through a time check, their system was able to reach 97% classification accuracy.

Body track consistency. using Earth movies, birds, and drones.

Rozantsev and associates. qualities of extracted photoflow [4]. CNN.

Classified the three categories with an average accuracy of 91% after being disaggregated.

When I mentioned these benefits in it. Sound is used to train the 1D CNN model.

Through obtaining drone sound samples and identifying the features of the Anti-Money Laundering Commission.

Giang et al.'s [5] use of acoustic characteristics allowed me to distinguish the drones from other bodies with 99% accuracy. Gyun and colleagues [6] assessed many dimensions. Using drone sound spectrometers, 1D and 2D CNN structure is created. CNN is a dual-dimensional network, they discovered.

It was more effective than previous models in differentiating between drones.
97% precision.

The author [7] offered a radio frequency-based method for identifying drones. (RF) Dispatched via Deep Learning Methodology (DL), namely the Telepathic Neural Network (CNN), during a live contact session between the drone and its controller.

Mejias and associates. I tracked the identification of unmanned small vehicles using morphological prepreparation channels and the hidden Markoff model [8].

METHODOLOGY

One kind of artificial neural network that is frequently utilized in computer visual and image analysis applications is the CNN. The architecture of CNN was created to capitalize on

The hierarchy's structure as it appears in the visual data.

Among the layers that make up a CNN are the entire conglomerate and its related layers. The rows are used to extract the features from the input image.

The filtering filters are applied in layers on it. to minimize computation and composition.

Due to grouping, the demands and spatial sizes of the maps are gradually decreasing rows.

Ultimately, the fully-connected layers classify and interpret the features.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-02 | June-2024

A CNN goes through two main phases of operation during training: feedforward and back propagation.

The input image is processed through the network layers and the output is computed during the feed forward stage. The weights and filters that make up the network are initialized at random. To update the parameters and reduce error, the output error is computed against the ground truth label and propagated backwards in the back propagation stage. Until the CNN can correctly classify images, this process is repeated numerous times.

CNNs are more efficient than fully-connected neural networks for computer vision tasks because of their architecture. To balance model accuracy and training time, careful hyperparameter tuning of the layer types, filter sizes, etc., is necessary. In general, CNNs are among the most widely used deep learning models in computer vision applications such as object detection, image recognition, and others .

The CNN consists of

- Input layer - 256x256x3 images
- 3 Convolutional layers:
- Layer 1 - 64 filters of 3x3, ReLU activation
- Layer 2 - 32 filters of 3x3, ReLU activation
- Layer 3 - 16 filters of 3x3, ReLU activation
- 3 Max pooling layers - 2x2 pool size
- Batch normalization and dropout after each convolutional layer
- Flattening layer
- 2 Fully connected layers:
- Layer 1 - 64 units, ReLU activation
- Layer 2 - 32 units, ReLU activation
- Output layer - 2 units with softmax activation

With the CNN Layer Architecture depicted in Figure .1.

Using the optimizer developed by Adam and categorical crossentropy loss, the model was trained over a period of 15 epochs.

After 15 training epochs, the model achieves a validation accuracy of 93%, according to the training logs. To increase accuracy, the number of epochs can be increased further until the validation loss stops decreasing. Overall, the three convolutional/pooling layers in this CNN architecture offer good results on the provided dataset. The layers are made to decrease the dimensionality of the input and gradually extract spatial features.

Input Layer (256x256x3 images)
Convolutional Layer 1 (64 3x3 filters)
Max Pooling Layer 1 (2x2 pool size)
Convolutional Layer 2 (32 3x3 filters)
Max Pooling Layer 2 (2x2 pool size)

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-02 | June-2024

Convolutional Layer 3 (16 3x3 filters)
Max Pooling Layer 3 (2x2 pool size)
Flattening Layer
Fully Connected Layer 1 (64 units)
Fully Connected Layer 2 (32 units)
Output Layer (2 units, softmax)

Table 1: CNN Layer Architecture

Dataset

We make use of an 828-image dataset from Kaggle that compares drones and birds. The 778 bird photos in the collection depict a range of species in various settings. There are 48 drone photos total, showing different consumer drone models from different perspectives and ranges. This disparity in class is a reflection of how hard it is to find diverse drone images as opposed to easily accessible bird datasets. Figure .2.

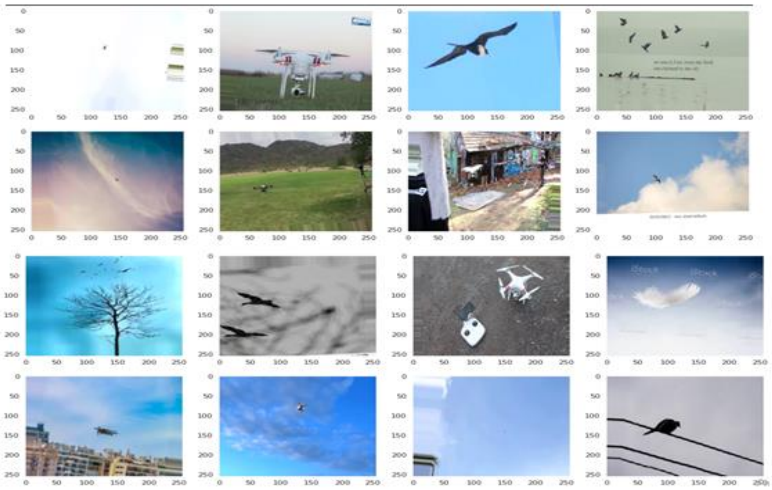


Figure 1: Images for Drone and Bird

Training Dataset:

- ✓ It contains 80% of the total data, which is approximately 662 images.
- ✓ These images are utilized for training the classification model.

Validation Dataset:

- ✓ It comprises the remaining 10% of the total data, equivalent to about 83 images.
- ✓ These images are used to assess the performance and accuracy of the trained model.

Testing Dataset:

- ✓ Also encompassing the remaining 10% of the total data, around 83 images.
- ✓ These images are employed to evaluate the performance and accuracy of the trained model after training.

This phase keeps a sufficient number of images for thorough testing and validation, while also making sure the model is successfully trained on a sizable dataset. It helps develop a reliable and accurate classification model that can effectively distinguish between images taken by drones and birds.

Table 2: Divide the Database into Training, Verification and Testing Data

Dataset Type	Number of Images	Percentage
Training Dataset	662	80%
Validation Dataset	83	10%
Testing Dataset	83	10%
Total	828	100%

Data pre-processing

Here are the image processing steps used in the provided code, organized and formatted : shown in the Figure .3.

Grayscale Conversion

- Using tf.image.rgb_to_grayscale

Resizing

- To256 x 256 pixels using target_size in ImageDataGenerator

Using Image DataGenerator:

- Random Rotation
- Up to 8 degrees using rotation_range
- Random Horizontal Shift
- Up to 0.1 using width_shift_range
- Random Vertical Shift
- Up to 0.1 using height_shift_range
- Random Zoom
- Up to 0.2 using zoom_range
- Random Horizontal Flipping
- Using horizontal_flip
- Random Brightness Adjustment
- Using brightness_range

Batching

- Using batch_size in flow_from_dataframe

Generating Training & Validation Sets

- Using validation_split in ImageDataGenerator

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-02 | June-2024

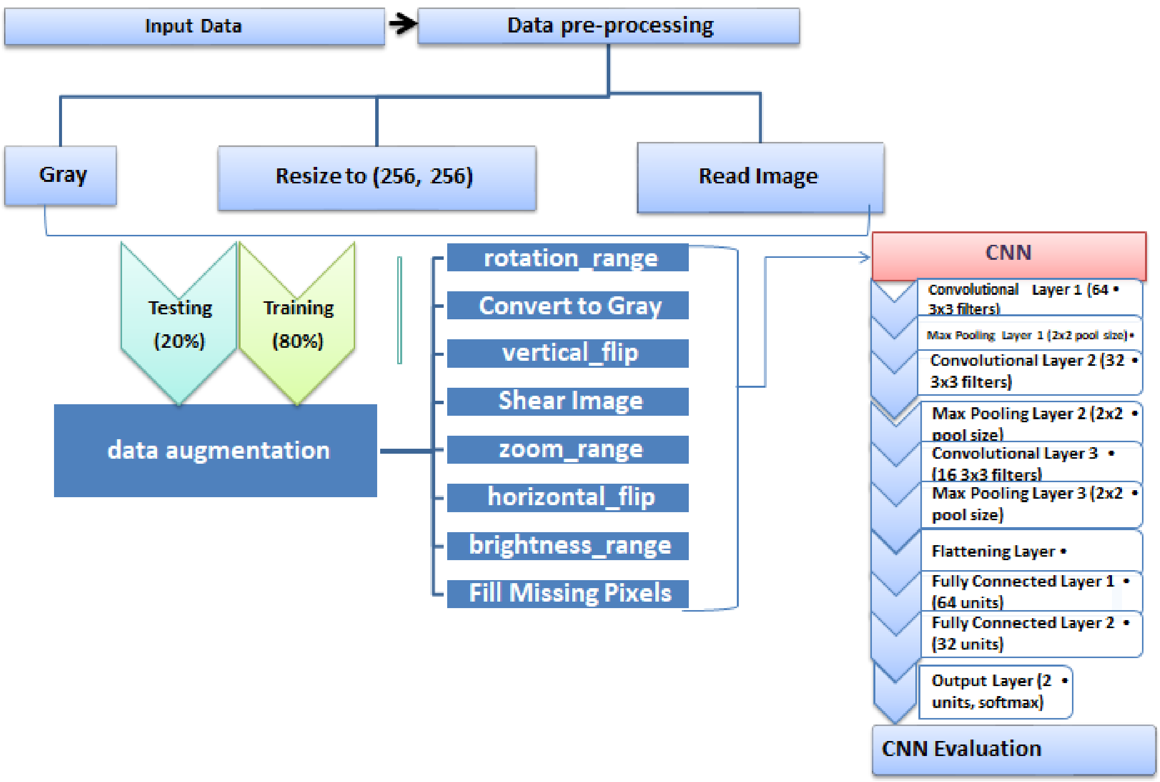


Figure 2: General graph for the Proposed Approach Methodology

RESULTS

Performance Evaluation Metrics

To evaluate the effectiveness of our model, we used multiple metrics specifically designed for the classification task. These metrics shed light on various facets of the accuracy and predictive power of the model.

Confusion Matrix

The confusion matrix is a square matrix of dimension (N N), where N is the number of classes. It is an essential tool in classification tasks. There would be two classes and a 2 by 2 matrix shown in the Figure 3. It contains the following values:

- True Positive (TP) cases are those that were appropriately identified as positive.
- True Negative (TN): Cases that were accurately classified as negative.
- False Positive (FP): Cases that were incorrectly identified as positive.
- False Negative (FN): Situations that were incorrectly predicted as negative.

Logical Creations Education Research Institute
LC INTERNATIONAL JOURNAL OF STEM
E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-02 | June-2024

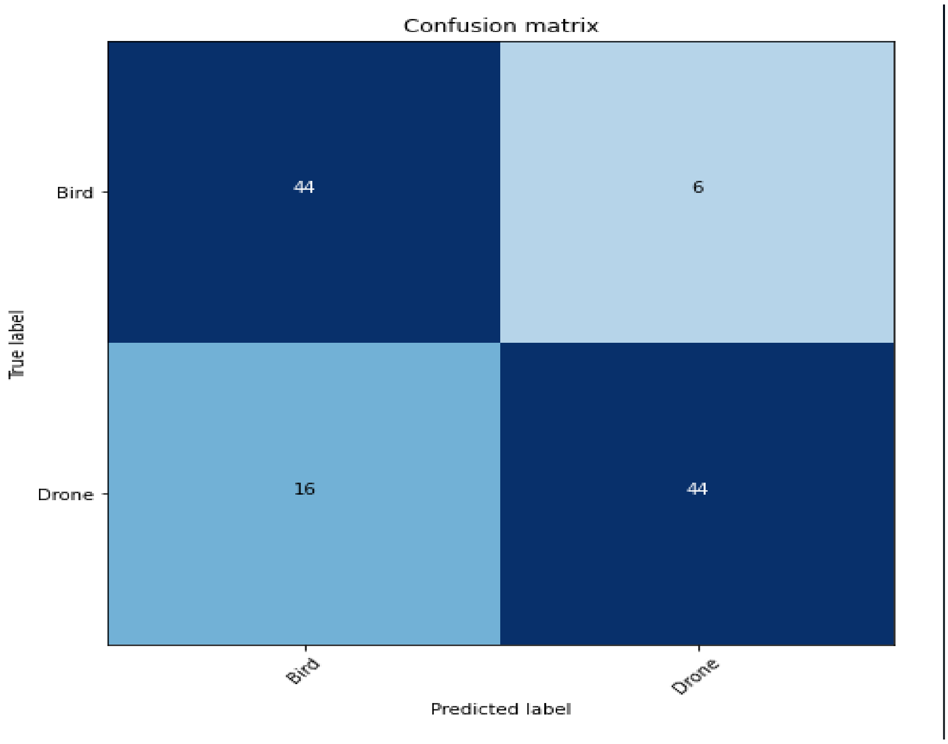


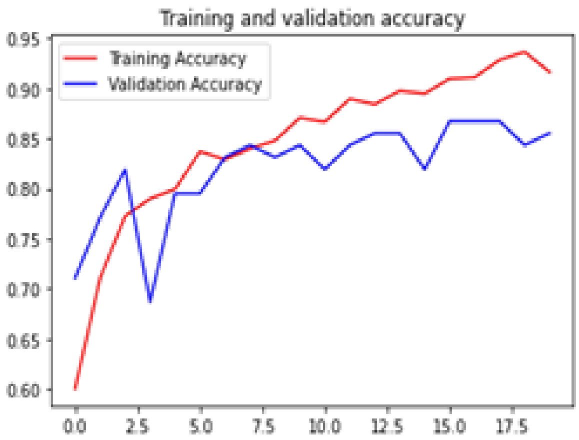
Figure 3: Confusion matrix extracted by the test phase

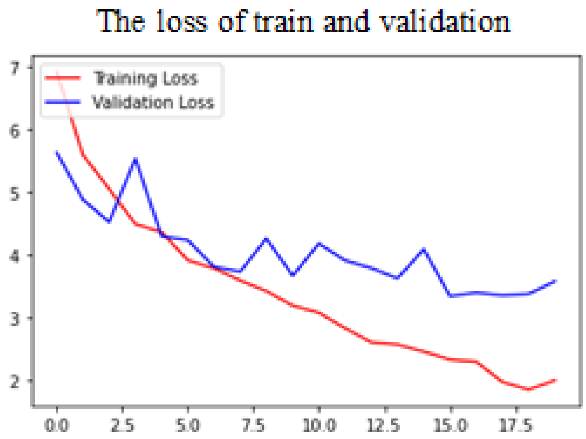
Accuracy

- Accuracy is defined as the proportion of all cases where the forecasts are accurate. Although it provides a broad measure of accuracy, imbalanced datasets may cause it to be misleading shown in the Figure .5.
- Accuracy = (TP + TN) / (TP + FP + FN + TN)

Table 3: Comparison of CNN results with results of other models

Model	Training Accuracy	Val Accuracy
Neural Network(NN)	73.1%	70.5%
Convolutional Neural Network(CNN)	93.2%	88.8%
Transfer Learning (ResNet50)	80%	79.6%





The system with which the above results are obtained are as follows:

- GHz Intel(R) Core i7-10510U CPU
- 2.30 GHz8 GB RAM
- Operating system 64-bit for Windows 10

Classification

Result of Drone Aircraft Classification for Birds shown in the Figure .7.

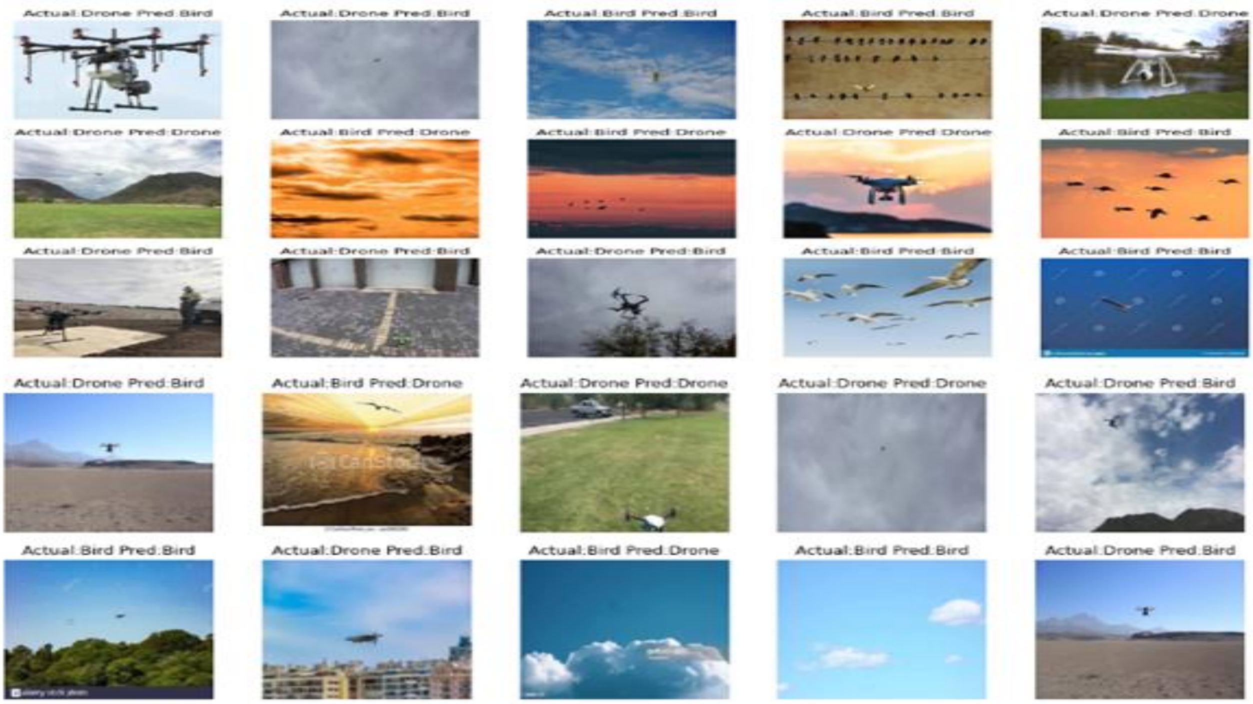


Figure 6: Drone Aircraft Classification for Birds

CONCLUSION AND RECOMMENDATIONS

Drones are now widely used for a variety of tasks in the modern world, including security, reconnaissance, and getting into restricted spaces. Because of their small size and low sound imprint, they are now considered essential. As a result, efficient drone detection has become crucial, especially when it comes to security.

The two-stage drone detection algorithm presented in this study consists of training and testing phases. During the training phase, a convolutional neural network (CNN) was trained using the carefully extracted features and attributes from the images. Later, during the experimental stage, the trained system showed an impressive level of skill in categorizing input images.

This method uses the CNN's strong classification capabilities to effectively handle the problem of differentiating between similar parts of drones and birds. The outcomes unmistakably show that the CNN model is better at distinguishing between the subtle features of these two objects than more conventional techniques like Neural Networks (NN) and Transfer Learning (ResNet50).

This research makes a substantial contribution to improving security in a variety of domains in addition to advancing the field of drone detection.

REFERENCES

[1] C. Lin, J. Luo, C. Chen, K. Chen, and C. Shen, "DRONE-BIRD: A Dataset and Method for Drone Detection in Videos Captured by Moving Cameras," IEEE International Conference on Multimedia and Expo (ICME), pp. 1-6, July 2019.

[2] A. Rozantsev, V. Lepetit, and P. Fua, "Detecting flying objects using a single moving camera," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 5, pp. 879-892, May 2017.

[3] W. Giang et al., "Flight Identification and Noise Monitoring System Using Low Cost Microcontroller and Convolutional Neural Network Model," Applied Sciences, vol. 10, no. 2, p. 559, January 2020.

[4] S. Jeon et al., "Empirical Study of Drone Sound Detection in Real-Life Environment with Deep Neural Networks," arXiv:1803.05531, March 2018.

[5] S. Al-Emadi and F. Al-Senaid, "Drone Detection Approach Based on Radio-Frequency Using Convolutional Neural Network," 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), Doha, Qatar, 2020, pp. 29-34, doi:10.1109/ICIoT48696.2020.9089489.

E-ISSN: 2708-7123
Web: www.lcjstem.com | Email: editor@lcjstem.com
Volume-05 | Issue-02 | June-2024

[6] L. Mejias, S. McNamara, J. Lai, and J. Ford, "Vision-based detection and tracking of aerial targets for UAV collision avoidance," in 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2010, pp. 87-92

[7] K. Mihreteab, M. Iwahashi, and M. Yamamoto, "Crow birds detection using HOG and CS-LBP," in 2012 International Symposium on Intelligent Signal Processing and Communications Systems, 2012, pp. 406-409.

[8] X. Li, Z. Huang, F. Wang, X. Wang, and T. Liu, "Toward Convolutional Neural Networks on Pulse Repetition Interval Modulation Recognition," IEEE Communications Letters, vol. 22, no. 11, pp. 2286-2289, 2018